Computers Math. Applic. Vol. 18, No. 1–3, pp. 127–131, 1989
Printed in Great Britain.
0097-4943/89 \$3.00 + 0.00

OPTIMAL STRATEGIES IN EVENT-CONSTRAINED DIFFERENTIAL GAMES

M. HEYMANN,[1] N. RAJAN[2] and M. ARDEMA[3]
[1]Technion, Israel Institute of Technology, Haifa, Israel
[2]Sterling Software, Palo Alto, CA 94043, U.S.A.
[3]Department of Mechanical Engineering, Santa Clara University, Santa Clara, CA 95053, U.S.A.

Abstract—Combat games are formulated as zero-sum differential games with unilateral event constraints. An interior penalty function approach is employed to approximate optimal strategies for the players. The method is very attractive computationally and possesses suitable approximation and convergence properties.

1. INTRODUCTION

In a recent series of papers [1–3], a mathematical formulation of combat games was proposed. By combat is meant a dynamical encounter in n dimensional space $X = \mathscr{R}^n$ between two players, or opponents, u and v, both of whom have offensive capabilities and objectives. The state transition is governed by a set of n ordinary differential equations

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(t, x, u, v), \tag{1}$$

with state $x = x(t) \in X$ and controls $u = u(t) \in U \subset \mathscr{R}^m$ and $v = v(t) \in V \subset \mathscr{R}^p$, where U and V are compact subsets. The initial time is t_0 and the initial state is $x(t_0) = x_0$.

Associated with the players are target sets $\mathscr{T}_u$ and $\mathscr{T}_v$ in *event space* $\mathscr{R} \times \mathscr{R}^n$, where $\mathscr{T}_u$ and $\mathscr{T}_v$ are closed and there exists a time $t^* > t_0$ such that $(t, x) \in \mathscr{T}_u \cap \mathscr{T}_v$ for all (t, x) such that $t \geqslant t^*$.

The game is said to terminate at time $\bar{t}$ where

$$\bar{t} := \inf\{t \geqslant t_0 \,|\, (t, x) \notin \operatorname{int}[\mathscr{T}_u \cup \mathscr{T}_v]\}. \tag{2}$$

Thus $\bar{t} \leqslant t^*$ and it is the objective of each player to terminate the game in his own target set while avoiding that of his opponent. If a player achieves this goal he is said to *win* the game. For each initial event the combat game can end with one of the following four outcomes: (a) player u wins, (b) player v wins, (c) a draw (the game terminates with $t = t^*$), and (d) joint capture (both players win with $\bar{t} < t^*$).

We shall adopt the concepts of strategy, value and saddle point as formulated in Friedman [4] with which the reader is assumed to be familiar. We shall say that a player's strategy is *winning* if it ensures him a win against all possible strategies of his opponent. Assuming that for a given event only one player can win (say u), u's optimal strategy is determined so as to minimize over all his winning strategies the cost functional $J = J_u$ given by

$$J = g(x(\bar{t}\,)] + \int_{t_0}^{\bar{t}} h[t, x(t), u(t), v(t)]\,\mathrm{d}t. \tag{3}$$

His opponent v chooses his optimal strategy so as to maximize J. The resulting differential game is a zero-sum game with unilateral event constraint $((t, x) \notin \operatorname{int}(\mathscr{T}_v) \forall t \in [t_0, \bar{t}]$ that u endeavors to satisfy and v to violate.

It will be assumed throughout the paper that the initial event is in u's winning zone, i.e. that u is the winning player and hence u is the minimizer and v the maximizer of the cost functional.

We shall make the following assumptions about the functions f, g and h:

Assumption 1. $f(t, x, u, v)$ satisfies suitable smoothness conditions that guarantee unique solution on $[t_0, t^*]$ for all measurable functions $u(t)$, $v(t)$ with values in U and V respectively.

Assumption 2. $h(t, x, u, v)$ is continuous in $[t_0, t^*] \times \mathscr{R}^n \times U \times V$.

Assumption 3. The functions f and h are separable in u and v, that is

$$f(t, x, u, v) = f_1(t, x, u) + f_2(t, x, v), \tag{4}$$

$$h(t, x, u, v) = h_1(t, x, u) + h_2(t, x, v). \tag{5}$$

Assumption 4. $g(x(t))$ is continuous on $\mathscr{R}^n$.

Assumption 5. For each (t, x),

$$\max_{v \in V} \min_{u \in U} h(t, x, u, v) > 0. \tag{6}$$

In the ensuing discussion we develop an interior penalty function approach to solve event constrained differential (combat) games.

2. EXISTENCE OF VALUE AND SADDLE POINT IN EVENT-CONSTRAINED DIFFERENTIAL GAMES

We derive in this section some preliminary results (in the spirit of Friedman [4]) on which our approach hinges.

Let (t, x) be any point of $\partial \mathscr{T}_v$ (the boundary of $\mathscr{T}_v$). We shall say that (t, x) is ***nonusable*** for v if there exists a $\bar{u} \in U$ such that for all $v \in V$ and all outward normals $\eta = \eta(t, x)$ to $\mathscr{T}_v$ *at* (t, x)

$$\eta^{\mathrm{T}} \cdot f(t, x, \bar{u}, v) \geqslant 0. \tag{7}$$

The point (t, x) is called *strictly nonusable* if the above inequality holds strictly.

We shall denote by Δ and Γ strategies of u and v, respectively. It is clear that if for a winning strategy Δ (of u) and any strategy Γ (of v), the corresponding trajectory contains points of $\partial \mathscr{T}_v$, they must be nonusable. In fact, we shall assume throughout the ensuing discussion that all such points of $\partial \mathscr{T}_v$ are actually strictly nonusable. We shall denote the set of all strictly nonusable points of $\partial \mathscr{T}_v$ by $\partial \mathscr{T}_v^s$. We introduce now the following parametrization.

Let $\mathscr{T}_v$ be given by a set of inequalities

$$\mathscr{T}_v := \{(t, x) \mid \alpha_i(t, x) \leqslant 0;\ i = 1, \ldots, I\}, \tag{8}$$

where $\alpha_i(t, x)$ are C^2 functions. The admissible region $\bar{X}_u$ of the game is then the union of the complement of $\mathscr{T}_v$ and of nonusable points of $\partial \mathscr{T}_v$.

For each $i = 1, \ldots, I$ let $\beta_i(t, x)$ be defined by

$$\beta_i(t, x) := [\alpha_i(t, x) + |\alpha_i(t, x)|]/2, \tag{9}$$

where $|\cdot|$ denotes the absolute value, and let

$$\beta(t, x) := \sum_{i=1}^{I} \beta_i(t, x). \tag{10}$$

Then $\beta(t, x)$ is nonnegative, and equals zero if and only if $(t, x) \in \mathscr{T}_v$.

We need the following additional assumptions:

Assumption 6. The strictly nonusable part $\partial \mathscr{T}_v^s$ of $\partial \mathscr{T}_v$ is C^2.

Assumption 7. The derivatives $\partial f_i(t, x, u, v)/\partial x_j$, $1 \leqslant i,j \leqslant n$ exist in some open neighborhood containing $\partial \mathscr{T}_v^s$.

Assumption 8. For every $(t, x) \in \partial \mathscr{T}_v^s$ there exists an open neighborhood $\mathscr{N}$ of (t, x) on $\partial \mathscr{T}_v^s$ with C^2 coordinates $t, \theta_1, \ldots, \theta_{n-1}$.

From the assumptions (A6) and (A8) it follows that for each point $(t, x) \in \mathscr{T}_v^s$, $(t, \theta, \beta) = (t, \theta_1, \ldots, \theta_{n-1}, \beta)$ forms a C^1 coordinate system in some X_u neighborhood M of (t, x). Therefore the differential system (1) can be written locally in M in the form

$$\frac{\mathrm{d}\theta_i}{\mathrm{d}t} = \hat{f}_i(t, \theta, \beta, u, v); \quad i = 1, \ldots, n-1$$

$$\frac{\mathrm{d}\beta}{\mathrm{d}t} = \hat{f}_n(t, \theta, \beta, u, v), \tag{11}$$

and for each $(t, x) \in M$,

$$\max_{v \in V} \min_{v \in U} \hat{f}_n(t, \theta, \beta, u, v) > 0. \tag{12}$$

Assumption 9. (1) There exists a positive number ϵ_0 such that for $i = 1, \ldots, n$ the partial derivatives

$$b_{ij}(t) := \begin{cases} \dfrac{\partial \hat{f}_i(t, \theta, \beta, u, v)}{\partial \theta_j}; & j = 1, \ldots, n-1, \\ \dfrac{\partial \hat{f}_i(t, \theta, \beta, u, v)}{\partial \beta}, & \end{cases} \tag{13}$$

exist and are continuous for all (t, θ, β) such that $0 \leqslant \beta \leqslant \epsilon_0$ and all control functions $u(t)$ and $v(t)$. (2) Consider now any pair of control functions $u(t)$ and $v(t)$ for which the corresponding trajectory satisfies $0 \leqslant \beta \leqslant \epsilon_0$ in some subinterval $(\bar{t}, \bar{\bar{t}})$ of (t_0, t^*) and let the $b_{ij}(t)$ be defined as in (13) for these functions. For $C > 0$ consider the linear system

$$\frac{\mathrm{d}z_i}{\mathrm{d}t} = \sum_{j=1}^{n} b_{ij}(t) z_j; \quad i = 1, \ldots, n; \; \bar{t} \leqslant t \leqslant \bar{\bar{t}}, \tag{14}$$

with

$$z_n(\bar{t}) = 1, \quad |z_i(\bar{t})| < C; \quad i = 1, \ldots, n-1; \tag{15}$$

then there exists a positive number θ_0 independent of $u(t)$ and $v(t)$, $\bar{t}$, $\bar{\bar{t}}$ (but dependent on C) such that

$$z_n(t) \geqslant \theta_0 \quad \text{if } \bar{t} \leqslant t \leqslant \bar{\bar{t}}. \tag{16}$$

Recall that a pair of strategies (Δ^*, Γ^*) is called an $\bar{X}_u$ *saddle point* if (i) the trajectory corresponding to (Δ^*, Γ) is winning for u for all Γ, (ii) the value exists and equals $J(\Delta^*, \Gamma^*)$ and (iii) for any admissible pair of strategies (Δ, Γ) the following inequalities hold:

$$J(\Delta^*, \Gamma) \leqslant J(\Delta^*, \Gamma^*) \leqslant J(\Delta, \Gamma^*). \tag{17}$$

The following central theorem can be proved:

Theorem 2.1

Assume that conditions (A1)–(A9) all hold and let (t_0, x_0) be in the interior of u's winning zone. Then the differential game associated with (1), (2), (3) and with event constraint $(t, x) \notin \mathrm{int}(\mathcal{T}_v)$, $t_0 \leqslant t \leqslant \bar{t}$, has both value and an $\bar{X}_u$ saddle point.

Theorem 2.1 depends among other things on the following lemma which we shall need independently below.

Lemma 2.2

Assume that conditions (A1)–(A9) all hold and let (t_0, x_0) be in the interior of u's winning zone. Then there exists $\epsilon^* > 0$ such that for each $0 \leqslant \epsilon \leqslant \epsilon^*$ there exists a winning strategy $\Delta(\epsilon)$ for u for which, given any strategy Γ of v, the corresponding trajectory satisfies the condition that $\beta(t, x) \geqslant \epsilon$ for all $t_0 \leqslant t \leqslant \bar{t}$.

In closing the present section we should like to emphasize that the existence of $\bar{X}_u$ saddle point strategies for the event constrained games does not imply that optimal *feedback strategies* exist. Indeed, the strategies computed via the Isaacs equation may be invalid on the boundary of the constraint set, a fact that can be illustrated by simple examples. To overcome the resulting difficulties we employ below a penalty function approach that, for almost all (t, x) in $\bar{X}_u$, yields arbitrarily close feedback approximations to the $\bar{X}_u$ saddle point strategies for the players.

3. INTERIOR PENALTY FUNCTION APPROACH

Let $\gamma:(0,\infty)\to\mathscr{R}^+$ be a C^1 function satisfying the conditions that:

Condition B1. $\gamma(\beta)$ is monotonically decreasing, i.e. $\mathrm{d}\gamma/\mathrm{d}\beta \leqslant 0$.
Condition B2.

$$\lim_{\beta\to 0}\gamma(\beta)=\infty.$$

Condition B3. For any $\beta_0>0$,

$$\lim_{\beta\to 0}\left|\int_{\beta_0}^{\beta}\gamma(\beta)\,\mathrm{d}\beta\right|=\infty.$$

Two examples of functions satisfying the above conditions are

$$\gamma(\beta)=\beta^{-1} \tag{18}$$

and

$$\gamma(\beta)=\begin{cases}\bar{\beta}\beta^{-1}+\beta\bar{\beta}^{-1}-2, & \text{for } 0<\beta\leqslant\bar{\beta},\\ 0 & \text{for } \bar{\beta}<\beta,\end{cases} \tag{19}$$

where $\bar{\beta}$ is any positive constant.

We define now the *penalty function*

$$J_p := \int_{t_0}^{\bar{t}}\mu(t,x)\,\mathrm{d}t \tag{20}$$

where $\mu(t,x) := \gamma(\beta(t,x))$. From condition (B3) it follows easily that if a trajectory approaches $\partial\mathscr{T}_v^s$ then J_p goes to infinity. More specifically the following theorem holds true:

Theorem 3.1

Let J_p be defined by (20) and assume conditions (A1) and (B3) hold. Then for every positive number $\mathscr{S}$ there exists a number ϵ such that if along a trajectory of (1) $\inf(\beta(t,x))\leqslant\epsilon$ then $J_p\geqslant\mathscr{S}$.

Let r be a positive number and consider now the cost functional

$$\tilde{J} := J + rJ_p \tag{21}$$

instead of the cost functional J of (3). It is readily noted that Theorem 2.1 remains true if (21) replaces (3) and, in fact, since J_p is unbounded, the saddle point of the game with payoff (21) will always yield an interior trajectory so that the event constraint can be removed. Thus, we formulate for r the *unconstrained* differential game defined by (1), (2), and (21) and obtain the following:

Theorem 3.2

Assume that conditions (A1)–(A9), (B1)–(B3) hold and let (t_0, x_0) be in the interior of u's winning zone. Then the (unconstrained) differential game associated with (1), (2) and (21) has value and an $\bar{X}_u$ saddle point as well as optimal feedback strategies almost everywhere.

Thus, for every positive number r the unconstrained "penalty game" defined above has optimal feedback strategies almost everywhere that can be computed by employing the Hamilton–Jacobi necessary conditions (Isaacs equation) for optimality (or for that matter, any other computational technique).

We shall show next that by suitable selection of the parameter r, arbitrarily close approximation of strategoes for the original constrained differential game will be obtained.

First we note that with (21) replacing (3), the essential objectives of the two playerse are still maintained. In particular, the minimizer u wishes to prevent constraint violation and hence is penalized if the trajectory gets too close to the constraint boundary. The maximizer, v, on the other hand, wishes to cause constraint violation if he can, and hence his objective is enhanced by the incentive given him through the penalty term.

Consider now a decreasing sequence of positive numbers $\{r_k\}$ with $r_k \to 0$ as $k \to \infty$. For each k, let Δ_k^* and Γ_k^* denote $\bar{X}_u$ saddle point strategies of the unconstrained penalty game (1), (2) and cost

$$\bar{J}_k := J + r_k J_p. \tag{22}$$

The following then holds:

Proposition 3.3

Let $r_{k+1} < r_k$. Then

$$\bar{J}_{k+1}(\Delta_{k+1}^*, \Gamma_{k+1}^*) < \bar{J}_k(\Delta_k^*, \Gamma_k^*).$$

From Proposition 3.3 it follows that the sequence

$$\{\bar{J}_k(\Delta_k^*, \Gamma_k^*)\}$$

is a monotonically decreasing sequence of positive numbers and hence must converge to a number $\bar{J}_\infty \geqslant 0$. We shall show below that $\bar{J}_\infty^* = J(\Delta^*, \Gamma^*)$ where $J(\Delta^*, \Gamma^*)$ is the value of the original event-constrainted differential game.

Theorem 3.4

Assume conditions (A1)–(A9), (B1)–(B3) hold and let (t_0, x_0) be in the interior of u's winning zone. For each r_k in $\{r_k\}$ consider the penalty game associated with (1), (2) and (22). Then for each $\epsilon > 0$ there exists a positive number K such that for all $k \geqslant K$,

$$J(\Delta^*, \Gamma^*) < \bar{J}_k(\Delta_k^*, \Gamma_k^*) \leqslant J(\Delta^*, \Gamma^*) + \epsilon. \tag{23}$$

The inequality (23) provides the essential justification for employing the proposed penalty function approach for actual computation of optimal approximating strategies for the two players.

4. CONCLUDING REMARKS

We have developed an interior penalty function approach for computation of (approximately) optimal strategies in (event constrained) combat games. Theorem 3.4 provides the theoretical justification for our approach.

It is interesting to make some qualitative observations.

In combat games, the winning player "initiates" the combat. If he chooses the saddle point strategy Δ^* (of the constrained game) then for every strategy Γ of his opponent,

$$J(\Delta^*, \Gamma) \leqslant J(\Delta^*, \Gamma^*), \tag{24}$$

so that his opponent has no incentive to deviate from the strategy Γ^*. If, on the other hand, u decides to play an optimal penalty strategy Δ_k^* (for any fixed chosen Γ_k), then it is true that

$$\bar{J}_k(\Delta_k^*, \Gamma) \leqslant \bar{J}_k(\Delta_k^*, \Gamma_k^*), \tag{25}$$

but it may still be true that

$$J(\Delta_k^*, \Gamma) > J(\Delta^*, \Gamma^*), \tag{26}$$

so that by selecting Δ_k^*, u is making a sacrifice in terms of cost (of the original game) in favor of "security". The inequality (23) however bounds this conceivable sacrifice to within arbitrarily small limits specifiable by the winning player u.

REFERENCES

1. M. D. Ardema, M. Heymann and N. Rajan, Combat games. *J. Optim. Theory Applic.* **46**, 391–398 (1985).
2. M. Heymann, M. D. Ardema and N. Rajan, A formulation and analysis of combat games. NASA Report TP 2487 (1985).
3. M. D. Ardema, M. Heymann and N. Rajan, Analysis of a combat problem: the Turret game. *J. Optim. Theory Applic.* **54**, 23–42 (1987).
4. A. Freidman, *Differential Games.* Wiley, New York (1971).